

0325.00344
CD00024

HYBRID DATA TRANSPORT SCHEME OVER OPTICAL NETWORKS

Cross Reference to Related Applications

This application claims the benefit of U.S. Provisional Application No. 60/184,264, filed February 23, 2000 and is hereby incorporated by reference in its entirety.

The present application may relate to U.S. Serial No. 09/523,576, filed March 10, 2000 (~~Attorney Docket No. 0325.00317~~), and Serial No. 09/523,476, filed March 10, 2000 (~~Attorney Docket No. 0325.00318~~) U.S. Serial No. 09/535,819, filed concurrently (~~Attorney Docket No. 0325.00345~~) and Serial No. 09/535,890, filed concurrently, (~~Attorney Docket No. 0325.00346~~) which are hereby incorporated by reference in their entirety.

Field of the Invention

The present invention relates to a method and/or architecture for hybrid data transportation generally and, more particularly, to sending a mix of different data types over a fiber optic network running SONET/SDH framing.

0325.00344
CD00024

Background of the Invention

5

Long distance and metropolitan area network (MAN) communications rely on short-haul and long haul fiber optic networks to transport data and telephony traffic. One conventional way to transmit data in fiber networks is through a Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) protocol. In a SONET/SDH network, data travels in fixed size envelopes that repeat every 125 microseconds. With this synchronous fixed-length framing, every byte (e.g., 8 bits of data) inside a SONET/SDH frame represents a 64 Kbps (64000 bits/sec) channel. The 64 Kbps channel has the same rate as supported by current telephone channels (also called DS0 channels).

10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
00

20

SONET was designed to efficiently carry telephony Plesiochronous Digital Hierarchy (PDH) channels such as T1/T3. This was easily achieved by dividing the payload area in fixed slots called virtual tributaries (VT). These virtual tributaries are then grouped together to form higher-rate channels. These fixed slots are efficient for carrying fixed-bandwidth telephony channels because any one or more channels can be added or removed from a bundle without processing an entire payload of channels. Because SONET frames repeat at fixed intervals, these virtual tributaries

0325.00344

CD00024

have fixed locations and time intervals, and it is easy to extract T1/T3 or fractions of T1 without processing the entire SONET payload.

With growing volume in data traffic, however, SONET/SDH networks must now carry a significantly large number of data packets - such as ATM (Asynchronous Transfer Mode - 53 bytes each) and IP (Internet Protocol - variable-size packets) in addition to traditional T1/T3 channels. The synchronous framing structure of SONET/SDH that is quite efficient for carrying T1/T3 channels is not able to carry both fixed-bandwidth and variable-bandwidth channels in an optimum way.

SONET/SDH has an inefficient utilization of fiber bandwidth for data packets. For data transport, some of the virtual tributaries ~~that~~ are created for transporting fixed-bandwidth T1 traffic while others are used for transporting packet data packets such as ATM and IP. Since an individual virtual tributary has a limited bandwidth, extra mechanisms have to be used for sending data packets of higher bandwidth using virtual tributaries.

In one technique, a 10Mbps data packet channel, for example, is inverse-multiplexed into smaller bandwidth streams and

0325.00344
CD00024

then sent on many virtual tributaries. At the other end, these streams are integrated to reconstruct the full 10Mbps channel. In another method, many of the virtual tributaries are concatenated using hardware to create a higher-bandwidth virtual tributary for 5 transmitting the high-bandwidth data packet.

SONET/SDH lacks of support for data mixing. A SONET fiber link carrying frames containing ATM cells cannot carry POS, because ATM cells frequently carry QoS-sensitive data such as CES (Circuit Emulation Service) or multimedia traffic. Introduction of SONET frames containing POS will cause significant delays (e.g., 125 μ s for each POS frame inserted in the link).

In each of these methods, a unique Path Signal Label (PSL) value in the POH (Path Over-Head) field of ^{the} SONET frame identifies the type of data transmission inside the payload. The payload area is also referred to as SPE (Synchronous Payload Envelope). Because a PSL value identifies contents of ^{the} entire SONET payload envelope, only one type of transmission can be supported at a time in a SONET frame.

One method for data transmission is to use the entire 20 SONET SPE for data packets. The SONET payload area is filled with IP packets using Packet-over-SONET (POS) packets. POS packets are

0325.00344

CD00024

packets by 0x7E (Hexadecimal) at both ends of a packet, with a framing using PPP (Point-to-Point Protocol). Many packets can be put inside a single SONET SPE. This method can only support variable-length packet protocol such as IP. A SONET fiber containing these packets cannot transport T1/T3 channels or real-time streams using ATM cells. The reason for this limitation is that each SONET SPE containing IP packets, for example, introduces a delay of 125 microseconds. Such a delay is not acceptable for T1/T3 circuits or real-time streams using ATM cells.

Another method for data transmission is to use the entire SONET SPE for ATM cells. In this case, SONET SPE is filled with ATM Cells. ATM cells are delimited by their fixed length, and are tracked by doing a hunt for their header checksum byte. Services such as T1, Frame Relay, Ethernet, etc. are transported over ATM using standard protocols. This requires complex implementations in hardware and incorporation of ATM service interworking at each service boundary.

Another method for data transmission is to use the virtual tributaries (VT) for data packets and ATM cells. In this method, a SONET SPE is partitioned in many fixed-bandwidth slots called virtual tributaries (VT). For data transport, some of these

0325.00344

CD00024

virtual tributaries may contain T1/T3 type of fixed-bandwidth traffic while others are used for transporting packet data packets such as ATM and IP.

Since an individual virtual tributary has a limited bandwidth, extra mechanisms have to be used for sending data packets of higher bandwidth using virtual tributaries. In one technique, a 10Mbps data packet channel, for example, is inverse-multiplexed into smaller bandwidth streams and then sent on many virtual tributaries. At the other end, these streams are integrated to reconstruct the full 10Mbps channel. In another method, many of the virtual tributaries are concatenated using hardware to create a higher-bandwidth virtual tributary for transmitting the high-bandwidth data packet.

Each of these methods uses a fixed-bandwidth channel or a set of channels for transmitting network data packets. In each method, bandwidth capacity of the fiber is poorly utilized since network data packets are bursty in nature and average bandwidth utilization is quite low.

Referring to FIG. 1, examples of various data types are shown. A set of time-division-multiplexed (TDM) packets 12a-12n, a set of ATM packets 14a-14n and a set of POS packets 16a-16n are

0325.00344

CD00024

shown in connection with a SONET fiber line 18. In a SONET network, only one type of data can be transferred at a time. The data is identified by a unique PSL (path signal label) byte value inside a Path Over-Head (POH) of the TDM packets 12a-12n,^{the} ATM 5 packets 14a-14n, the POS packets 16a-16n, of PDH traffic. The nodes at different points in the SONET fiber line 18 have different types of data to send on the network.

Referring to FIG. 2, various data transmissions are shown in a SONET ring 30. Assume, for example, that a node A has ATM cells 14a-14n to transmit on the SONET ring 30. The SONET ring 30 forms a SONET synchronous payload envelope (SPE), sets the PSL value for ATM cells (in the POH) and sends the SONET frame down a link 32a-32n. Even if the SPE is only partially full, the entire SPE frame size is transmitted. If a node B has IP packets to send, the node cannot use the partially filled SONET frame received from the node A to add POS packets, because the PSL value identifies only one type of data (ATM cells, in this example). The node must now wait 125 μ s for the partially filled packet to transfer down the link 32b-32n. Furthermore, the SONET ring 30 cannot handle T1 or 20 T3 channels entirely in a payload space. The only way it can simultaneously support T1/T3 channels is by using virtual

0325.00344

CD00024

tributaries and by using some of the tributaries for non-T1/T3 traffic.

Data transmission on Ethernet, ATM, Frame Relay, PPP, and other network types use different data packet formats for each 5 transmission type. Each transmission type implements some type of error detection technique. The error detection lets nodes detect if packets are received with data integrity. CRC (Cyclic Redundancy Check) is used for (i) LAN (Local Area Network) applications using Ethernet and (ii) WAN (Wide Area Network) 10 applications using PPP (Point-to-Point Protocol).

The CRC is the result of a polynomial division on preceding bytes (in this case user data, layer2 and layer3 addresses). The sending station (e.g., transmitting node) computes the CRC on bytes being transmitted and then appends the CRC value at the end of data transmission.

A receiving station (e.g., receiving node) receives all 15 bytes in sequence and computes the CRC value based on the received byte stream. The receiving station then compares an independently calculated CRC with the CRC value received from ^{the} sending station. 20 If the two CRC values do not match, the receiving station discards the packet.

0325.00344

CD00024

With Ethernet networks, user data is encapsulated with layer 2 (data link) and layer 3 (network layer) addresses as headers, followed by a 16/32-byte Cyclic Redundancy Check as shown by the following TABLE 1:

5

TABLE 1

Layer 2 Addresses		Data Identifier	Layer 3 addresses	User Data	Error Detection
Destination MAC (6 bytes)	Source MAC (6 bytes)	Protocol Identifier or IEEE802.3 Length Field (2 bytes)	...Network Layer addresses...	Payload	CRC (Cyclic Redundancy Check)

An example of PPP frame format is defined by the following TABLE 2:

TABLE 2

Frame Boundary	Address	Control Word	Layer 3 Identifier, etc.	User Data	Error Detection	Frame Boundary
0111	Address	Control	PPP Network ID	Payload...	CRC	0111 1110

PPP is used in point-to-point wide area network applications such as dial-up modems, frame relay node-to-node communications or any communication where a link establishment is needed before actual transmission can take place.

15

20

0325.00344

CD00024

Referring to FIG. 3, a system 40 illustrating an example of a packet moving across a number layer 2 bridging devices 42a-42n is shown. The packet moves from a node 44 (e.g., A) and a node 46 (e.g., B). The packet is modified during transit, but not by the layer 2 devices 42a-42n. In a local area network, devices such as layer 2 transparent bridges usually do not modify the packets (such as Ethernet packets). The bridge devices 42a-42n do not recompute CRC as the packet travels from one bridge device (e.g., 42a) to another bridge device (e.g., 42n). Each bridge device 42a-42n checks the CRC for correctness, but does not recompute the CRC value on an outgoing port of each bridge device 42a-42n. Recomputation is not necessary since the content of the packet does not change.

Referring to FIG. 4, a system 60 illustrating packet modifications for a number of layer 3 routers and MPLS switch devices 62a-62n is shown. Layer 3 routing and MPLS can result in packet modification. When a layer 3 device 62a-62n (such as a router or a layer 3 switch) route data, the device 62a-62n usually modifies layer 2 and/or layer 3 addresses as the packet 62a-62n crosses through ports of the devices 62a-62n. Such modification

0325.00344
CD00024

causes the devices 62a-62n to recompute the CRC on the packet being transferred.

MPLS (Multi-Protocol Label Switching) is a technique where, particular routes that a packet travels are assigned a series of labels (e.g., 20 bits each out of a 32-bit word). The labels are inserted between layer 2 and layer 3 address values (for Ethernet frames), and before network layer values (for PPP frames). The labels are changed, added or dropped as the packet travels from one node 64 (e.g., A) to another node 66 (e.g., B). Because the contents of the packet change, the nodes A and B recompute the CRC before retransmitting the packet on a respective output port.

Referring to FIG. 5, a system 70 illustrating an example of a problem in error detection due to CRC regeneration in error detection when CRC is recomputed is shown. The system 70 is illustrated in the context of a SONET network. Generally, a node that receives data before another node is called upstream node, while node that receives data sent by the upstream node is called downstream node. A node (e.g., A) sends a frame down a link 72. If some bit errors occur at an upstream node (e.g., B, C and/or D) that receives a packet with a correct CRC, the downstream node (A) will never learn about the bit errors if the upstream node (B, C

0325.00344

CD00024

and/or D) recomputes the CRC for the packet to be sent before transmission.

When MPLS is used, a node that receives the transferred packet usually (i) swaps the label with a different value, (ii) 5 pops the label, and/or (iii) adds a new label to the stack. If MPLS is embedded inside the packet, the payload CRC will change at each node (e.g., A, B, C and D). When an intermediate node has internal memory or other device errors that result in data corruption (e.g., at node B), or has a logic error that shifts incorrect data out on the output port, the node B recomputes CRC on the outgoing packet and sends the frame with the recomputed CRC value. In such an example, the downstream node (e.g., node C) will find the packet intact, since the CRC received from the upstream neighbor was computed on modified bytes. If the node (C) is not the final destination, the error will propagate and will not be detected by down stream nodes (e.g., node D).

When an application program on the destination node (e.g., node D) gets the packet and discards the packet, the destination node (D) does not know which of the upstream nodes (A, 20 B and/or C) caused the error. The destination node (D) may not

0325.00344

CD00024

even know there was an error if the error occurs in any of the non-critical data areas of the packet.

Various conventional protocols have been developed to improve bandwidth usage by attempting to partially solve two problems in SONET networking (i) lack of support for data mixing, and (ii) bandwidth reuse limitations. Such conventional protocols have been partially able to achieve additional bandwidth either by creating fatter pipes with VT, by filling payloads with ATM cells carrying different types of data, or by using SONET link as a multi-node access network using Ethernet framing for data packets.

While virtual tributaries provide an efficient way to transport PDH traffic, allocation of bandwidth for LAN data such as 10/100Mbps using a SONET fiber containing T1 traffic becomes difficult. To support 10Mbps through Virtual tributaries one has to either dedicate an entire STS-1 (51Mbps) frame, or use several virtual tributary channels and to perform inverse multiplexing.

Virtual concatenation concatenates several VT channels into bigger virtual pipes to carry higher bandwidth traffic. With such a protocol, while some virtual tributaries carry T1/T3 data as usual, others are concatenated for transport of higher bandwidth data traffic such as 10/100Mbps links.

0325.00344

CD00024

5

However, virtual concatenation allocates a fixed bandwidth for LAN. Virtual tributaries cannot dynamically adjust bandwidth usage on a packet-by-packet basis. While it is possible to change the concatenated bandwidth through software, such an implementation does not yield much for bandwidth utilization. Bursty LAN traffic bandwidth usage is typically quite low, resulting in significant waste when used over a fixed bandwidth virtual pipe (average use of a 10Mbps link is 20%, and if an entire STS-1 is used the efficiency becomes about 4%). Another problem with virtual concatenation is that while one virtual pipe may be overloaded with traffic, others may be underutilized. Virtual concatenation cannot dynamically adjust network loads on different channels.

Another conventional approach to increase the utilization of SONET bandwidth is to fill the payload area with ATM cells using a technique known as ATM VP (Virtual Path) multiplexing. ATM VP multiplexing encodes packets in ATM cells and then inserts the cells inside a SONET SPE. The ATM VP multiplexer typically utilizes CES to carry DS0/1 PDH traffic.

20

However, operations, administration and maintenance (OAM) operations of ATM network are different from that of SONET, and

0325.00344

CD00024

management of the two protocols can become difficult. Similarly, ATM network routing and switch-to-switch signaling data paths are different from IP network routes, resulting in network operational complexity.

5 Network traffic statistics monitoring has shown that a significant percentage (about 45%) of network traffic comprises IP packets with small packet sizes (e.g., around 40 bytes). With IP over ATM (rfc2684 - Multiprotocol Encapsulation over ATM AAL5) payload size slightly exceeds what can fit inside a single ATM cell. Such an arrangement results in transmission of two ATM cells, with the second cell that is mostly ATM overhead and stuffing bytes. This, with other ATM overheads, means having to allocate extra SONET bandwidth for IP transport if ATM is used as a transport protocol.

100-200-300-400-500-600-700-800-900-1000

20

In addition, using ATM for sending different services requires implementation of ATM transport protocols and interworking for all related protocols (such as IP-over-ATM, Frame Relay-ATM Internetworking, circuit emulation, etc.) in the device. Such implementation requires a high level of complexity in hardware and software, resulting in higher costs of manufacturing and operation of networks.

0325.00344

CD00024

Such conventional network architectures are not efficient for transmitting variable-length IP packets that form the majority of data network traffic on the Internet. Data network traffic comprises variable-length IP packets and fixed-length ATM cells
5 (i.e., 53 bytes).

As more and more data is being transported on SONET/SDH rings, there is a need to send variable-length packets on pre-existing SONET/SDH networks. These packets originate out of routers and other data access devices. While SONET/SDH networks must transport these data packets, they must also continue to support TDM-style fixed length packets for telephony and leased line applications.

Conventional approaches cannot mix POS with ATM cells in a single SONET SPE. Conventional approaches do not leave some area of the SONET SPE reserved for VTs and others for POS and/or ATM cells.

Limitations of conventional approaches include one or more of the following (i) inability to mix TDM channels with packet-oriented data over SONET/SDH rings due to timing constraints
20 of the TDM channels without a fixed bandwidth virtual tributary mechanism and (ii) limiting data channels sent on fiber carrying T1

0325.00344

CD00024

lines to Virtual tributaries (since virtual tributaries are of fixed bandwidth, this restriction limits data channels to fixed bandwidth operation).

Therefore, if a SONET SPE is carrying all ATM cells, it
5 cannot carry IP variable-length packets, and vice versa. Such switches route all packets coming at one tributary to another until switching paths are changed through reprogramming.

Conventional approaches do not take into account network loading conditions and do not support dynamic bandwidth provisioning. Conventional approaches also have increased traffic on core links from too much concentration into too few links. Transporting IP traffic over virtual tributary channels that was originally designed for DS0/1/3 connectivity is inefficient. Because VT assignment is fixed, IP transport is not able to take advantage of total available SONET/SDH bandwidth.

In conventional approaches, IP packets are constrained to go through some pre-configured VT channels while other VT channels may be under-utilized. Once a VT channel is dedicated for a particular traffic and is put on a specific circuit-switched path,
20 the topology does not change, even if traffic conditions change.

Conventional approaches have the following disadvantages

- (i) ATM and Packets are implemented on different rings because of QoS and timing issues; (ii) very high cost for new fiber and SONET equipment for Telco/ISP/MAN; (iii) only one type of packet goes inside a SONET SPE at one time (the remaining bytes of frame on SONET are wasted) (SONET packets go around the entire SONET ring, limiting bandwidth); and/or (iv) the only way to support telephony channels along with data packets is to allocated part of SONET frame for packet data transmissions (which results in an inefficient bandwidth usage).

Summary of the Invention

One aspect of the present invention concerns a packet configured to transmit information via a network. The packet may comprise a header error portion and a payload error portion.

Another aspect of the present invention concerns an apparatus comprising one or more nodes coupled to a network. Each node may be configured to receive and/or transmit one or more of a plurality of packets. Each packet may comprise a first portion configured to store payload error information and a second portion configured to store header error information.

0325.00344

CD00024

The objects, features and advantages of the present invention include providing a method and/or architecture that may (i) provide a single SONET ring to carry a number of forms of data; (ii) allow a number of forms of data types to operate inside a single SONET SPE; (iii) provide bandwidth reuse at various nodes of a network; (iv) quickly detect and possibly bypass a faulty node on a network, (v) provide extremely easy fault analysis in medium to long-haul optical networks, (vi) improve packet transport performance, and/or (vii) provide savings in network administration and maintenance expenses for optical networks.

Brief Description of the Drawings

These and other objects, features and advantages of the present invention will be apparent from the following detailed description and the appended claims and drawings in which:

FIG. 1 is a detailed block diagram of conventional fiber optics transmission data types;

FIG. 2 is a block diagram illustrating a mixing operation of a conventional SONET network;

FIG. 3 is a block diagram of a conventional Ethernet packet moving across a number of layer 2 bridged devices;

0325.00344
CD00024

FIG. 4 is a block diagram of a conventional Ethernet packet moving across a number of layer 3 routers or MPLS switch;

FIG. 5 is a block diagram of a conventional SONET ring network;

5 FIG. 6 is a block diagram of a preferred embodiment of the present invention;

FIG. 7 is a more detailed block diagram of the circuit of FIG. 4;

FIG. 8 is a block diagram of a HDT frame inside a SONET SPE;

FIG. 9 is a detailed block diagram of a variable length envelope;

FIG. 10 is a detailed block diagram of a packet of FIG. 8;

FIG. 11 is a more detailed block diagram of a header of FIG. 9;

FIG. 12 is a flow chart illustrating an operation of the present invention;

20 FIG. 13 is a flow chart illustrating an operation of the present invention;

0325.00344
CD00024

FIG. 14 is a flow chart illustrating an operation of the present invention; and

FIG. 15 is a block diagram illustrating an operation of the present invention.

5

Detailed Description of the Preferred Embodiments

The present invention may provide a Hybrid Data Transport (HDT) protocol that may allow transmission of fixed bandwidth channels (e.g., T1/T3), variable-bandwidth data sources (e.g., ATM), IP and any other protocol data in a single SONET frame using a single fiber network. The protocol of the present invention may work seamlessly across a mix of SONET and non-SONET networks, and may yield cost savings in fiber infrastructure, equipment, and operation.

With the use of the Hybrid Data Transport (HDT) protocol of the present invention, an existing fiber network may be fully utilized to transport a number of different types of traffic. The present invention may additionally dynamically manage bandwidth usage on a packet-by-packet basis.

20 The present invention may provide spatial reuse of bandwidth, allocation of PDH bandwidth in 64Kbps increments,

0325.00344

CD00024

protocol-independent MPLS (Multi-Protocol Label Switching) support, and/or seamless operation over point-to-point and ring networks with SONET/SDH, direct data-over-fiber configurations or other network configurations. The present invention may also be
5 applicable to Synchronous Digital Hierarchy (SDH). However, SONET is used as a general description for SONET/SDH networks with similar implementation for SDH networks. Additional details of the operation of the HDT protocol are also described in connection with direct data-over-fiber networks.

Referring to FIG. 6, a block diagram of a system 100 is shown in accordance with a preferred embodiment of the present invention. A more detailed implementation of the system 100 is illustrated in FIG. 7. The system 100 may comprise a number of devices 102a-102n connected to a network backbone 104. FIG. 7 illustrates the addition of a number of chips 106a-106n. The devices 102a-102n may receive T1/T3 signals, ATM signals, and POS signals. The devices 102a-102n may receive data in one or more of the following data transmission media: SONET, SDH, direct data over fiber (e.g., both in point-to-point or ring configuration) with or
20 without SONET/SDH framing and other transmission methods needed to meet the design criteria of a particular implementation.

The system 100 may provide an increase in the data traffic handling capabilities of SONET/SDH networks. The system 100 may implement a design of a SONET/SDH add/drop multiplexer (ADM) (and a SONET/SDH cross-connect) that may function on variable-length packets. With such an approach, IP (or other protocol) packets of different lengths may be (i) added to a SONET/SDH SPE and (ii) terminated at a different network node (e.g., devices 102a-102n).

Referring to FIG. 8, a detailed block diagram of SONET/SDH payload envelope (SPE) 200 is shown. The present invention may embed a header (and/or footer) 202 (e.g., a 32-bit packet header) to create a deterministic packet transport protocol. The packet header may comprise a 32-bit payload header 204a-204n that may precede each frame, regardless of the particular packet type stored within the frame. The protocol identification may be implemented as a few header bits configured to denote the particular type of packet (e.g., ATM, IP, PPP, Frame Relay, etc.) embedded within the payload portion of a particular frame. Bandwidth maximization may be implemented with another bit in the header 202 that may specify whether the packet may be reused by the intermediate SONET nodes 102a-102n. The SONET framing may be left

0325.00344

CD00024

unchanged by implementing a single PSL (Path Signal Label) value 206 in a SONET Path Over Head (POH) 208 that is generally able to specify the various types of packets embedded within the payload of a particular frame. The system 100 may be directly applicable to 5 WDM/DWDM Fiber because individual packet framing is independent of SONET. The system 100 may be also used in IP-over-Fiber networks.

The present invention may use a header CRC separate from a payload CRC. A separate CRC for the header bytes may change from node to node and may leave the payload CRC unchanged. Particular nodes may (possibly) modify header bytes such as MPLS (Multi-Protocol Label Switching) label values, IP addresses, and any other parameters that change from node to node. Such nodes may recompute the CRC only for those particular bytes. However, while a particular node generally checks the payload CRC upon input for packet correctness, the particular node does not recompute payload CRC when the packet is transmitted on an output port.

TABLE 3 may illustrate an Ethernet Frame Data Format. On Ethernet networks, user data is generally encapsulated with layer 2 (data link) and layer 3 (network layer) addresses as headers, 20 followed by a 16/32-byte Cyclic Redundancy Check as shown by the following TABLE 3:

0325.00344

CD00024

TABLE 3

Layer 2 Addresses		Data	Layer 3 addresses	Header CRC	User	Error
Destination MAC (6 bytes)	Source MAC (6 bytes)	Identifier Protocol or IEEE802.3 Length Field (2 bytes)	Network Layer addresses	Field Header CRC	Data Payload	Detection CRC
5						

TABLE 3 may illustrate an example of a PPP frame format.

A header CRC value is generally inserted before user data starts as shown in the following TABLE 4:

TABLE 4

Frame Boundary	Address	Control Word	Layer 3 Identifier, etc.	Header Field	User	Error Detection	Frame Boundary
01111110	Address	Control	PPP Network ID	Header Field CRC	Payload	CRC	01111110

An example of an Ethernet Frame Data Format with MPLS may be shown in TABLE 5. The header CRC field is generally inserted after the MPLS labels, and before the user data.

TABLE 5

0325.00344

CD00024

Layer 2 Addresses		Data Identifier	MPLS Labels	Layer 3 addresses	Header CRC Field	User Data	Error Detection
Destination MAC (6 bytes)	Source MAC (6 bytes)	Protocol Identifier or IEEE802.3 Length Field (2 bytes)	One or more 32-bit words	...Network Layer addresses...	Header CRC	Payload	CRC

5

TABLE 6 may illustrate a PPP frame format with MPLS. A header CRC value is generally inserted after the MPLS labels, and before user data starts as follows:

TABLE 6

Frame Boundary	Address	Control Word	PPP Identifier, etc.	MPLS Labels	Header CRC Field	User Data	Error Detection	Frame Boundary
01111110	Address	Control	PPP Network ID	One or more 32-bit words	Header CRC	Payload	CRC	01111110

DRAFT - 2/24/98

The CRC system described may be used for transmission over Optical Networks. In one example, the HDT frame header format

15 may be implemented with the CRC system described. A header CRC is generally introduced just before the actual payload starts.

The present invention may provide extremely easy fault analysis in medium to long-haul optical networks. The present invention may provide the ability to immediately trap data errors 20 to avoid allowing such errors to propagate from node to node.

0325.00344
CD00024

Savings in network administration and maintenance expenses for optical networks may result. Improved packet transport performance may also result since bad packets do not travel undetected on the SONET network. Detecting errors early may avoid retransmissions 5 and recovery at higher network layers.

The CRC system may allow device to quickly detect and possibly bypass a faulty node on a network. In one example, in optical fiber rings, a downstream node may identify a faulty upstream neighbor and proceed to perform a node or ring bypass operation to avoid the node.

The SONET/SDH payload envelope 200 (e.g., transmitted every 125μS) divided into variable length packets is shown. The header 202 may comprise one or more of the following parameters:
(i) packet length, (ii) length of CRC (Cyclic Redundancy Check),
(iii) payload identifier header to describe the nature of packet,
(iv) route labels that may help route packet inside network, (v)
payload header CRC, (vi) actual payload, and/or (vii) payloads CRC.

In SONET, a basic unit of transmission is a Synchronous Transport Signal Level 1 (STS-1) or Optical Carrier Level 1 (OC1) 20 signal. Both operate at 51.84 Mbit/s. STS-1 describes electrical signals, and OC1 refers to the same traffic after being converted

0325.00344

CD00024

into optical signals. SONET also allows channels to be multiplexed. An OC12 circuit, for instance, may carry traffic from four OC3 links. An OC12 circuit may also carry a single channel, in which case the line is said to be concatenated. Such circuits
5 may be described as OC3c, OC12c, and so on.

Another protocol similar to SONET is Synchronous Digital Hierarchy (SDH), defined by the ITU (International Telecommunication Union) as G.707 (shortly after ANSI formally ratified the T.105 spec for SONET). Although interconnection of SONET and SDH networks may be relatively rare, several new transoceanic telecommunications projects make use of such links.

The system 100 may maximize fiber bandwidth by implementing the Hybrid Data Transport (HDT) Protocol. The HDT protocol may allow dynamic management of packets to maximize bandwidth. The system 100 may allow the transport of different types of data over a single fiber link. With the system 100, IP (or other protocol) packets, Packet-Over-SONET (POS), ATM cells, G.702-based PDH (T1/T3), SRP, Frame Relay, and other types of data may be mixed inside a SONET payload and dynamically and sent on a
20 single fiber (as shown in FIGS. 5 and 6). The system 100 may provide robust scrambling and unified packet transport over ring

0325.00344

CD00024

and point-to-point networks and may be well suited for non-SONET configurations such as point-to-point WDM networks. The SONET SPE 200 may be filled with HDT frames that may carry a wide mix of fixed and variable bandwidth data. The Simple Data Link (SDL) framing protocol prefixes a payload with a 32-bit word, 16 bits of which hold the length of the packet and the other 16 bits contain CRC (Cyclic Redundancy Check) for the length field. SDL may provide a robust CRC-16 based framed boundary delineation mechanism compared to hybrid mix of point-to-point and ring topologies.

The system 100 may implement a single fiber link that may be used for sending different kinds of traffic to use the full capacity of the link. The system 100 may allow VT or sub-VT channels to be eliminated. ATM cells, IP (and other protocols) packets, PPP, frame relay, NxDS0, T1/T3 and others may be mixed inside SPE on a packet-by-packet basis.

PDH channels (such as T1/^{T3}/E1) may be dynamically allocated anywhere inside SONET payload in 64Kbps bandwidth increments. Bandwidth may be reusable in fine granularity in 64Kbps increments, with any type of data. For example, an IP packet may be dropped at 20 a node (e.g., B) where the node B may reuse the packet area for

0325.00344
CD00024

inserting ATM cells, frame relay, PDH traffic, or any other data type.

The system 100 may transport many packets of one or more different data types. The particular type of data may be placed 5 inside a single SONET SPE or data-over-fiber frame while preserving time dependency of data packets, such as PDH. The system 100 may be implemented without terminating the whole link capacity at each node. Destination nodes may indicate a start of the packet inside a SONET payload. Other packets may pass through the node.

10 Direct data-over-fiber configurations (e.g., without SONET framing) may be easily supported with full link monitoring and management. Support may also be provided for variable-size 15 packet SONET add/drop multiplexer (SONET ADM) devices. Variable-size packets may be transported inside a SONET SPE and nodes may cross-connect and add/drop the packets on different ports.

20 The system 100 may provide protocol-independent transport of MPLS labels. The HDT may provide for transmission of MPLS labels outside of protocol frame (rather than embedding such labels between data link and network layers as in conventional approaches). Intermediate SONET and optical fiber nodes may be

0325.00344
CD00024

used to create switching and add/drop systems without having to become protocol-aware.

With provision for MPLS labels transported outside of packets, the HDT may allow setting up of fast-reroute 5 fail-over paths over a hybrid of ring and point-to-point networks using MPLS labels, without requiring nodes to be protocol-sensitive. Such an implementation may eliminate the need to provision and reserve completely unused backup paths on SONET rings.

In operation, the payload headers 204a-204n may precede every packet to carry the information to support HDT. A uniform structure of the header across a variety of packet types generally simplifies design of optical nodes for packet processing for both SONET and direct data-over-fiber networks. The headers 204a-204n may also contain a reusability bit that is set by the sending node. If the reusability bit is cleared, a destination node may reset the data identification bits to free up the packet area for reuse by new data packet. The packet area may be reused at either at the destination node or other downstream nodes.

20 By using the payload header 204a-204n to identify the type of a packet, the HDT is generally able to extend data

0325.00344

CD00024

identification beyond PSL-based SPE-level data typing and put multiple data types inside a SONET SPE. A value of 0000 may indicate the packet area (e.g., the length of the packet area is given by the length value in the outside SDL framing) that does not generally contain any useful data and can be reused for storing new data. HDT may easily support traditional PDH and other guaranteed bandwidth channels. In SONET networks, a frame repeats every 125μS, resulting in a bandwidth of 64KHz for every byte in the payload. By fixing the starting location of some packets inside the SONET frame, slots may be created for sending TDM-style traffic. Because the packet length is changeable in one-byte increments, such slots may be created in increments of 64KHz. Because packets are dynamically created, fixed bandwidth channels may be created on the fly by clearing the reusability bit in the payload header.

Referring to FIG. 9, a detailed block diagram of the SONET SPE 200 is shown. The SONET SPE 200 may comprise a number of packets 220a-220n and a number of empty packets 222a-222n. The packet payload header 204a of the packet 220a may identify the packet/protocol. The packet payload header 204a may identify a packet type of the packet 220a stored (or transported). The

0325.00344

CD00024

payload header 204a may tell what kind of packet/protocol (such as Ethernet, PPP, IP, Frame Relay, ATM cells, T1, etc.) is inside a payload of the packet 220a. Different protocols may be supported at two ends (e.g., the devices 102a-102n) of a network without the
5 need for provisioning in advance. In contrast, conventional approaches use a protocol over WAN which is usually negotiated between two parties at the ends devices 102a-102n of the WAN link.

The payload header 204a may be used to tell whether one or more of the empty packets 222a-222n inside the SONET SPE 200 may be reused at an intermediate node. In contrast, in conventional SONET networks, the entire SONET SPE 200 travels around the ring until removed by the sender. With the system 100, a receiver may mark the SONET SPE 200 as reusable. Nodes on the fiber network 100 may mark different sections of the SONET SPE 200 as reusable by the other nodes 102a-102n.

Provisioning of TDM channels may provide the ability to mark a portion (or many portions) of a SONET SPE payload area as non-reusable. With a non-reusable area, even when a receiver receives the packet, another receiver cannot reuse the packet area.
20 However, the same receiver may reuse the non-reusable area.

In general, there is no limit to the order and manner of packet positioning. Any packet may be marked in any fashion to support, for example, a dynamic mix of data and voice (TDM) traffic on a SONET/SDH network. Such an implementation is not possible
5 with current technologies. The present invention may solve the problem of mixed value and data transmissions faced by telephone carriers and data providers.

As SONET frames containing fixed bandwidth channels move around the ring, intermediate nodes may detect these packets (e.g., the reusability bit is reset), note the offsets of these packets, and preserve the respective offsets when recreating the frame (e.g., after adding packets from local input ports) for outbound traffic.

Referring to FIG. 10, a detailed example of a packet is shown. An SDL framing 262 may be in the first 16 bits and may contain the length of the entire payload, including SDL framing bytes. A 16 bits of CRC-16 264 may be provided on the length field (e.g., $x^{16} + x^{12} + x^5 + 1$). The payload header 204a-204n may be a 32-bit word, followed optionally by an OAM bytes or MPLS labels
20 268. MPLS/OAM bytes may be variable number of MPLS labels or OAM values that may be transmitted in the header area of HDT, outside

0325.00344

CD00024

of payload. A next fragment offset 270 may be a 16-bit value showing the location offset of next packet fragment (if any) of the packet. The next fragment offset 270 is generally taken from the start of current packet. A header CRC 272 may be computed over 5 payload header bytes only, with same scrambling polynomial used for SDL framing. A payload area 274 may contain the actual packet to be transmitted over the WDM or SONET link. The payload area 274 may contain one of a number of types of protocol packets, such as Ethernet, ATM, GR.702, PPP, Frame Relay, etc. A payload CRC 276 may be user-controlled value and may be computed for the payload bytes only. The payload CRC 276 is generally either a 16 or 32-bit value, depending on mutual negotiation between sending and receiving stations.

Referring to FIG. 11, various parameters of the packet header 204a are shown. The particular bit width of the payload header 204a may be varied accordingly to meet the design criteria of a particular implementation. A packet identifier 280 (e.g., D3:D0) generally identifies the type of packet in the payload. For example, value of 0000 may represent a null packet. A null packet 20 may indicate that the payload area may be reused. When a packet is

0325.00344

CD00024

dropped at a node, the length field does not generally need to be modified for the packet, only the D3: D0 bits need to be cleared.

A header data area 282 may carry MPLS labels (e.g., outside of payload area). Operation administration and maintenance (OAM) bytes 282 may be used for link management, or any other data separately from the payload. A reusability area 284 (e.g., D7) may be a "1". If a SONET node can reuse a particular packet area, the size of the packet area may be given by the packet length field 264 of the SDL header. If the bit D7 is set to a "0", then a node will not generally mark the packet area as re-usable, even after a packet has been dropped. The particular nodes of the various configuration bits may be varied (e.g., inverted) accordingly to meet the design criteria of a particular implementation.

A header length area 286 (e.g., D15: D8) may include, in one example, a 32-bit payload header. A fragment identifier area 288 (e.g., D17: D16) may be implemented as a two word value. A value of 00 may indicate that the payload area contains a complete packet. A value of 01 may indicate the beginning packet of a fragmentation sequence. A value of 10 may indicate a continuation of packets. A value of 11 may mark the last fragment in the series. Other particular bit patterns may be implemented

0325.00344

CD00024

accordingly to meet the design criteria of a particular implementation.

A padding area 290 (e.g., D18: D19) may indicate a minimum packet length. In one example, the minimum packet length 5 may be 4 bytes (e.g., 2 bytes length + 2 bytes CRC). Idle bytes at the end of packets and elsewhere may be marked by a length field of "0000". In instances there may be less than 4 bytes left between packets. In this case, it may be impossible to place a SDL null packet. Such idle bytes are shown as tail-end padding for the preceding packet. An unused area 292 (e.g., D31:20) may be used for additional expansion.

Referring to FIG. 12, an example of a receive operation is shown. A node may receive a frame at a block 300. A block 302 may determine if the received frame is an HDT frame. The block 302 may use the PSL value in the POH to determine the type of protocol carried inside the SPE. If the PSL shows POS, ATM, or PDH traffic, the receive operation may proceed to the block 304. If no HDT packets are present, a block 306 handles the POS/ATM/PDH packet. If in the block 302, the PSL shows the SPE contains HDT frames, the 20 node uses additional logic for HDT processing to detect and route different types of packets embedded in the SONET SPE 200.

0325.00344

CD00024

5

A block 308 may read the POH. A block 310 may determine a first packet of the SONET SPE 200. A block 312 may read a length and CRC of the first packet. A block 314 may determine a match of the length and the CRC. If a non-match of the length and CRC occurs, the receive operation is generally set to a block 316. The block 316 may read a next word of the packet from the SONET SPE 318. If a match occurs, the receive operation may process the packet. Once the payload header has been processed and different packet types are been processed and different packet types are identified, hardware (e.g., implemented in the system 100) may use header fields to retrieve the payload and use usual hardware blocks for processing.

10
11
12
13
14
15
16
17
18
19
20

20

ATM cells are generally retrieved by first looking at the PSL value to determine their presence and then reaching the SONET SPE to get fixed byte ATM cells, either with or without HEC-based cell delineation. For example, if the payload header in the HDT shows the payload contains ATM cells, the hardware device generally retrieves payload bytes (up to number of bytes specified in length field) and sends the byte stream to an existing ATM cell processing logic. The ATM cell processing logic may then work on the byte

0325.00344

CD00024

stream using HEC hunting just as if the SPE contained only ATM cells in its payload.

Referring to FIG. 13, an example of a processing operation 320 is shown. A device supporting Hybrid Data Transport HDT protocol generally operates much the same as a normal SONET/SDH transport operates. Operations for processing ATM cells, POS, and PDH protocols are the same and illustrated as processing blocks 350, 352 and 354. HDT adds a header to packets to allow their mixing within the same SPE 200. Much of the HDT processing is generally related to processing of the header to identify the type of packet and then passing the starting address of data bytes to standard logic for handling the individual packet type. Support of PDH-type channel typically requires a fixed starting location for the channel in every frame. If PDH support is not needed, packets of any mix may be put anywhere inside the SPE 200 to achieve excellent bandwidth utilization without much operational complexity. When fixed bandwidth channels are carried, some data packets may need to be fragmented when the packet hit a static location. Fragmentation of a packet, however, is generally easily achieved in SONET networking because all bytes in the SPE 200 are

0325.00344

CD00024

transmitted sequentially. Additionally, recovering fragments and putting the fragment together may be simply accomplished.

Referring to FIG. 14, an example of a transmit operation 400 is shown. A device supporting HDT may receive a packet to be transmitted from a system side. In the transmit operation 400, a node may take inputs from different sources 402, encapsulate the packets with an SDL length/CRC fields 404, add an HDT header 406 to each of the packets, and then store the packets inside the SPE. The node may not send a fresh frame on the network in order to transmit the packets. A TDM channel check 408 may determine a reusability of the SPE 200. The transmit operation 400 may reuse available space in an incoming SPE (containing HDT frames). The transmit operation 400 may then proceed to a length check 410 to see if there is any space available to insert the packet to be sent. If there is enough space, the entire packet is stored (with proper SDL framing and HDT header bytes). Any remaining bytes, depending on the size, are generally either (i) filled with a null HDT packet (e.g., the payload header identification bits are 0000), (ii) filled with SDL null packets (e.g., pairs of length/CRC with a null length field), or (iii) accounted for as tail-end padding (e.g., if the size is less than 4 bytes).

0325.00344

CD00024

If the transmit operation 400 runs into a fixed-bandwidth channel allocation midway through the packet allocation, the packet is generally fragmented. In this case, a portion of the packet may be stored at one place and other fragments may be stored at another 5 free location. The first fragment offset pointer may contain the starting location of second fragment. Because bytes are transmitted sequentially in the SPE 200, reassembling fragments may be easily achieved.

If a particular node detects an incoming SONET frame on a receive port, or if there is a frame in the transmit/receive queue, the node checks the frame to see if there are unused/reusable areas in the incoming/queued frame that can be used for sending data. If there is enough space available in the frame, the node fills the space with additional data before sending the frame out.

In HDT, PDH channels of any bandwidth (up to allowable SONET bandwidth limits) may be provisioned anywhere inside the SONET SPE. To achieve precise timing, PDH bytes must begin at the same offset inside the SONET SPE. However, allocation of PDH 20 channels at different locations inside a SONET SPE may create fragments of unused bytes all over the SONET SPE. For efficient

0325.00344
CD00024

transport of variable-size IP packets, these unused bytes may be utilized for IP data.

Referring to FIG. 15, an example of spatial reuse with HDT is shown. Spatial reuse of bandwidth across a number of network nodes (e.g., A, B, C, D) may be achieved by permitting full or partial termination of individual packets at any node. Spatial reuse of bandwidth reclaimed from the terminated packet may increase performance. HDT may provide an ideal way to achieve spatial reuse of SONET bandwidth. Using add/drop of hybrid data, nodes can reuse released bandwidth for transmission of any of the various kinds of data.

0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

As the SONET SPEs are received at the nodes A, B, C, D, initial bytes may be placed in a small transit buffer. Through MPLS labels contained in the header section or through internal packet fields, a particular node A, B, C, D may be able to determine whether the packet belongs to the node. If the packet does not belong to the node, the bytes are streamed out of transit buffer to the output port. However, the packet may belong to the node A, B, C, D if, for example, (i) the D7 bit is set in the payload header, (ii) the packet area has been reserved for a fixed bandwidth channel such as a PDH, and/or (iii) in this case, the D3:

0325.00344
CD00024

D0 bits are not cleared. If the D7 bit is set to "0", the node may clear the D3: D0 field to mark the packets void and reusable, where the bytes belonging to the packet are sent to system. The number of bytes sent to the system may be specified in the length field of 5 the SDL header. If the header shows fragmentation then a packet is received in many fragments and sent out to the system until the last fragment is received.

Packets may be added either using a fresh SONET SPE or by reusing bytes inside an incoming or previously queued frame. The decision of which packet to add to a void or reusable packet area inside an SPE can be made on following lines by (a) selecting a packet (or a collection of packets) that will fit inside the reusable area, (b) selecting all packets that can fit inside the reusable space, or (c) selecting a packet based on QoS parameters or packet priority. Since SONET frames repeat at 125ms intervals, packet transmission may be arranged to achieve a desired rate. Once a packet is selected for addition to the SPE, the node creates a payload header by setting payload type, reusability and other bits for the packet. The circuit 100 may then add the header to 20 the payload.

0325.00344

CD00024

5

The SDL framing mechanism may use length/CRC pair information as a header and a frame delimiter. SDL provides a robust scrambling and frame locator technique and may be used for direct data-over-fiber networks where SONET framing may not be used. Implementing OAM packets may eliminate the need for complex SONET framing and link management overheads. In the example of direct data-over-fiber networks, the HDT protocol structure may operate unmodified. Point-to-point WDM networks and ring-based SONET networks (or any other network) may easily be mixed and connected to each other. With a powerful support for MPLS (that may be transported independently of payload), networks may be designed that may have alternative LSP (Label Switched Paths) links for a highly robust redundancy. For example, nodes on a SONET ring may be connected through another network that may be entirely different from the ring. The backup path could be a high-speed point-to-point link or a ring network that may be geographically quite diverse. By providing a common network protocol engine, the HDT protocol may permit configuration of these networks quite easily without requiring complex protocol translation logic for 20 different network configuration.

0325.00344

CD00024

The present invention may use a packet payload header to identify the kind of packet inside. These identifier bits tell what kind of packet/protocol (e.g., such as Ethernet, PPP, IP, Frame Relay, ATM cells, T1, etc.) is inside the payload. Using 5 such a technique, different protocols may be supported at two ends without the need for advanced provisioning. Using conventional methods, the use of a single protocol over a WAN needs to be negotiated.

0
10
20
30
40
50
60
70
80
90
100

The identifier may indicate whether one or more packet areas inside SONET SPE may be reused at an intermediate node. Conventional SONET networks require the SONET frame to travel around the ring until removed by the sender. Even when the receiving node received a packet, the frame went around the network, wasting bandwidth. With the present invention, not only may a receiver mark a SONET SPE as reusable, but different receivers on the fiber network may mark different sections of SONET SPE area as reusable when a packet is received by different receivers.

The present invention may provide the ability to mark a portion (or many portions) of a SONET SPE payload area as non-reusable. With such an implementation, when a receiver 20

0325.00344

CD00024

5

receives the packet, the packet area is not generally reused by another receiver/transmitter. However, the same receiver may reuse the marked payload area for add/drop applications. Allowing the same receiver to re-use a packet may help TDM channels and packet data within a single SPE.

Over time, bit definitions inside a payload header may change as further research is conducted on the fiber data protocol operation. Such changes in bit definitions are common in data communication protocols and do not change the nature and content of present invention.

The present invention often refers specifically to SONET. However, Synchronous Digital Hierarchy (SDH) protocols are equally appropriate. SDH is similar to SONET with differences in bit framing. These framing differences, however, do not change the discussion and scope of the present invention.

20

While the invention has been particularly shown and described with reference to the preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made without departing from the spirit and scope of the invention.